International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Novel ML Approaches for Treasury Forecasting: A Literature Survey

Atharv Joshi

Student, St. Francis High School, Mountain View

Abstract
U.S. government bonds are affected by central bank decisions. Bonds are less easily traded than stocks and the public data about them are not abundantly avail- able. In this project, we review the state-of-the-art methods in machine learning (ML) and artificial intelligence (AI) methods employed in forecasting interest rates for U.S. treasuries of varying maturities. Our work will highlight how powerful AI techniques can be leveraged in more accurate predictions of movement in treasury/government bonds.

1. Introduction

U.S. government bonds are investments that play a crucial role in the economy. The treasury securities finance government operations and they are used by the Federal Reserve to implement monetary policy and serve as reserve assets on the global scale. Treasury yields, specifically 10-year treasury yields, are important indicators of U.S. economic health and investor sentiment. They tend to rise during periods of expected economic expansion and inflation and fall during economic uncertainty or recession. Additionally, the relationship between short-term and long-term yields, known as the yield curve, shows upcoming economic downturns with an inverted curve. Increasing yields usually suggest strong economic growth for the future and often rising inflation expectations; decreasing yields suggest economic slowdown or investor uncertainty.

Central banks make decisions that affect bond markets; their influence on interest rates directly impacts bond yields and prices. The relationship between interest rates, bond prices, and bond yields is inversely correlated.

Treasury prediction involves econometric models coupled with time series analysis. However, financial markets are highly complex. These datasets include but are not limited to historical yield curves across various maturities, macroeconomic indicators (such as GDP growth, unemployment rates, and inflation), market sentiment data, trading volumes, and more. At the same time, computing power has improved significantly, leading to more advanced techniques. Machine Learning (ML) has been established as a valuable tool for financial predictions, though there are classical models with better predictive power (e.g., ARIMA). By using these large datasets and identifying complex patterns, both ML and deep learning (DL) algorithms may forecast trends in U.S. treasury bonds [1].

In this literature review, we study new ML approaches used in predicting U.S. treasury rates of various maturities. We examine forecasting methods, ML, and deep learning (DL) techniques, with a special emphasis on Transformer models, with their application to the treasuries data. Additionally, we discussed the impact of state-of-the-art (SOTA) in AI in forecasting.

2. Traditional Forecasting Methods

Classical statistical techniques have been used to forecast future values based on historical. Armstrong *et al.* reveal that typically, these methods consist of time series analysis, regression models, and classical decomposition techniques of time series data such as trend, seasonality, and noise components 2. In general, traditional forecasting approaches use exponential smoothing, and autoregressive integrated moving average (ARIMA) models are often used by classical approaches to capture linear relationships in data, efficiently laying the foundation for forecasting.

Classical models for analyzing financial markets, such as ARIMA, have clear limitations in capturing the complex dynamics present in real-world market behavior, as they are often not able to incorporate the non-linear relationships that are apparent in real financial market behavior. For instance, the average ARIMA error was 30% greater during the 2008 financial market crisis than when the market was more or less stable 3. This shows how limited ARIMA's scope is of accounting for non-linear financial dynamics and shocks.

These traditional models face several limitations 3 :

1. Linear assumptions for a non-linear world: ARIMA and similar models assume that financial data follows linear patterns, which may not be the case, as it is affected by various factors like investor behavior and unexpected events.
2. Ineffectiveness during extreme events: Classical models also assume a stable environment in the economy. This would weaken their forecasting abilities during sudden shifts in the market.
3. Parameter selection uncertainty: Implementing models like ARIMA requires choosing appropriate values for various parameters, which can make the analysis subjective and increases uncertainty.
4. Ineffectiveness for long-term forecasting: ARIMA models are best in forecasting short to intermediate-term forecasting. These models tend to perform poorly on long and complex time-series data as they fail to analyze structural changes and long-term market dynamics. For example, ARIMA models predicting stock prices five years ahead had an average error exceeding 50%.
5. Exclusion of external factors: ARIMA models consider only the past time series data and not macroeconomic or geopolitical factors. For this reason, finance professionals usually enhance ARIMA with other models, such as Vector Autoregression (VAR), to take into account such factors or variables. The use of ARIMA should be complemented with other techniques for predicting market behavior being impacted by changes in growth or inflation.

These limitations underscore the need for more sophisticated approaches that can better capture the complex dynamics of financial markets and provide more accurate forecasts and insights.

3. Machine Learning Approaches

Compared to the traditional methods, new machine learning techniques have improved modern treasury forecasting. In this section, we reviewed various machine learning approaches that can be divided into two different groups: supervised and unsupervised learning techniques.

3.1 Supervised Learning Techniques

Puglia and Tucker studied several supervised learning methods for forecasting U.S. recessions using treasury term spreads and other financial market and macroeconomic variables 4. One of the most popular supervised learning approaches to classification and regression problems is Support Vector Machine (SVM). SVMs are useful in treasury forecasting, as they can be used to handle high-dimensional data and find complex patterns in smaller datasets; most importantly, patterns indicating future economic downturns or recessions. Because SVMs can handle high-dimensional data and perform well on binary classification and regression problems, they can be employed to predict for treasury yields, economic indicators, and other continuous outputs in financial time-series data.

Random Forest is a stable, ensemble greedy algorithm that can be used for classification and regression, that bags multiple decision trees to improve prediction accuracy and reduce overfitting. Given that ARIMA is mostly suited for linear time series, Random Forest can capture linear and non-linear patterns, which makes it better for treasury forecasting. Furthermore, Random Forest multiple input features together, allowing for a better analysis of different economic indicators. It also reveals the influence of the variables on the forecasts made which is not easily available in ARIMA models. Random Forest is more robust to outliers and can manage missing data. It aggregates predictions from multiple decision trees, making it more reliable than single model methods such as ARIMA. Puglia and Tucker employed Random Forecast in forecasting recession and showed that though more traditional methods, such as probit regression, have been widely used in this area, machine learning methods like random forests can still provide competitive performance.

Gradient boosting methods, including XGBoost, are advanced ensemble techniques that build models sequentially, with each new model attempting to correct the errors of the previous ones 5. XGBoost owes its popularity in the field of financial forecasting and economics to its accuracy and run-time efficiencies 6. The results of Puglia and Tucker's paper indicate that although methods based on trees seem to outperform the probit regression model according to k-fold cross-validation, more conservative validation strategies still show that in some cases, probit regression is preferable, especially when using Nested Time Series (NTS) cross-validation method (that avoids data peeking).

3.2 Unsupervised Learning Techniques

Martin and Poczos highlighted the effectiveness of several unsupervised learning techniques for financial forecasting 7. There have been many studies that group stocks and bonds using clustering methods, such as Principal Component Analysis (PCA), Isomap, and hierarchical clustering, based on price patterns. This type of market segmentation is valuable in forecasting since models can find similarities and differences between groups of financial instruments.

4. Deep Learning Approaches

4.1 Advanced Neural Network Architecture

An early study of the adaptation of novel AI techniques was performed by Cheng *et al.* 8. In their studies, Cheng *et al.* employed neural networks (NN) to predict the direction of U.S. Treasury Bonds on a weekly basis. In five years, their NN model had an accuracy rate of approximately 67% for buying decisions and an annual return rate of 17.3%. Although it performs well, during market changes this model faces issues and the process of training the large networks was computational heavy. The study laid the groundwork for future applications of neural network architecture in finance, especially treasury bond research.

In more recent efforts, Sako *et al.* analyzed the use of both Recurrent Neural Networks (RNNs) and their variant Long short-term memory (LSTMs) in financial time series forecasting 9. RNNs are seen as a very useful tool, particularly in financial data analysis such as treasury yields. RNNs keep track of its previous inputs, and this renders the neural network cell to detect temporal patterns more efficiently. This feature makes RNNs well-suited for forecasting tasks where past information is critical in predicting future values. Long Short-Term Memory (LSTM) network is one of the variants of RNNs. Their main goal is to combat the vanishing gradient problem faced by conventional RNNs. By employing a much more complex structure involving memory cells and gating mechanisms, LSTMs can retain information longer, thus making them optimal for predicting yields on treasuries which are influenced by various factors over time.

Financial time series can also be modeled using Convolutional Neural Networks (CNNs), primarily because of their ability to learn local patterns and short-term dependencies 10. CNNs can be adapted to process sequential data by applying one-dimensional CNN kernels that convolute over the time series. This allows the model to identify short-term patterns in time series, rendering them useful for trends observed over a short time span. However, CNNs are unable to learn long-term dependencies due to the limited receptive field. In financial forecasting, where predicting accurately requires knowledge about both short-term movements and long-term trends, this limitation becomes a problem. To solve this problem, recent studies have proposed hybrid models that combine CNNs with Transformers 10. This hybrid approach combines the advantages of both architectures: CNNs for short-term pattern recognition and Transformers for capturing long-term dependencies. Such integration of CNNs with Transformers can predict intraday stock price changes, as well as other financial activities. Nonetheless, for the hybrid model to be more effective and suitable for real-world financial forecasting scenarios, there should be further research on CNNs as there are many intricacies in financial data.

4.2 Transformer Models

Autoformer is a deep learning architecture for long-term series forecasting 11. It uses the inner decomposition block that switches between decomposition and refinement. To capture complex temporal patterns, the Autoformer structure separates long-term trends from hidden variables. The available autocorrelation can help in the discovery of periodicity and aggregation of similar sub-series leading to its robustness for capturing long-range relationships. While autocorrelation is also fundamental to ARIMA models, Autoformer's approach is superior to it and modeling using Transformers for several reasons: its ability to handle non-linear patterns and non-stationary time series, to automatically learn complex features from the data, and to process longer sequences and capture intricate dependencies. This is useful particularly in treasury forecasting since long-term trends are important to analyze; this is what makes the CNN and Transformer based time series modeling (CTTS) so effective. While the encoder focuses on modeling seasonal parts by providing past seasonal information for cross-information in refining predictions, the decoder improves the prediction by using the past seasonal information respectively. Among these benchmarks are six different ones where each of them is considered within a long-term setting: energy, traffic, economics, weather, and disease with a 38% relative improvement of Autoformer. As prediction lengths increase it becomes clear that Autoformer's performance improves, thus besides treasury forecasting, the model can be easily generalized to other forecasting problems such as weather forecasting and consumption scheduling.

4.2.1 Challenges and Limitations of Transformer Models

However, other research has cast doubt on Transformers as the best way to forecast long- term time series. Zeng *et al.'s* study on Transformers provides a comprehensive evaluation that is evidenced against the utility and of Transformer-based time-series models 12. The authors used a very simple deep-learning algorithm consisting of only two linear layers of neurons, referred to as LTSF-Linear, which outperformed complicated attention-based forecasting mechanisms and models by a large margin (measured by MSE, Mean Squared Error) on various data types (e.g., traffic and weather).

A major concern with Transformer models is their strong propensity for overfitting. Due to their intricate architectures and vast number of parameters, Transformers can easily learn noise in the training data rather than the underlying patterns. This over- training can lead to false predictions, where the model identifies spurious correlations that do not exist in the real world. In contrast, classical statistical models, which often rely on simpler structures, are less prone to such pitfalls. They take into consideration fundamental relationships in the data, thus being more applicable in situations where deep learning techniques are not accommodative of complex dynamics.

The study also discusses the disadvantages of overfitted Transformer models that stem from the fact that treasury markets are complex and depend on factors such as economic indicators, interest rates and other financial variables. A model that overfits historical data may not be able to generalize future conditions. On the other hand, classical models are simple and more understandable and understand such patterns, since they are generally built based on domain knowledge and economic theory.

Furthermore, when applied in real-time forecasting scenarios, these Transformer models have certain limitations in terms of computational complexity, particularly due to their quadratic time and memory requirements which arise from the self-attention mechanism.

This mechanism calculates attention scores for all pair of elements in the input sequence. Thus, its time complexity is $O(L^2)$ where L refers to length, which increases the computational requirements quite rapidly as the length of the sequence increases. However, in treasury markets, it seems as if the efficiency of simpler models like LTSF-Linear may offer a practical advantage. LTSF-Linear algorithm has a time complexity of $O(L)$, linear with respect to sequence length, and an $O(1)$ maximum signal traversing path length, which allows it to capture both short-range and long-range temporal relations effectively 12. The processing of large amounts of data is crucial in treasury markets, making simpler models like LTSF-Linear more practical due to its forecasting performance.

5. Conclusion

The field of treasury forecasting has significantly changed due to the adoption of innovative machine learning methodologies: traditional, machine learning, deep learning methods, as well as the use of Transformers. Traditional forecasting methods, such as ARIMA, have demonstrated limitations, particularly during periods of market volatility, with an average error increase of 30% during the 2008 financial crisis 3. Deep learning models, which include RNNs and LSTM networks have recently been successful in overcoming the temporal dependencies that are observed in financial time series data. Autoformer is listed here as an example of a transformer architecture that represents the forefront of long-term forecasting capabilities. However, many obstacles are yet to be tackled in this space, such as overfitting risks and computational complexity with more advanced models. In the future, the potential for AI in treasury forecasting is most likely seen by hybrid modeling, combining different strengths to interpret complicated models better.

References

1. Estrella, A., & Mishkin, F. S. (1995, December). *Predicting U.S. Recessions: Financial variables as leading indicators* (Working Paper No. 5379). National Bureau of Economic Research. https://doi.org/10.3386/w5379
2. Armstrong, J. S. (2006). Findings from evidence-based forecasting: Methods for reducing forecast error. *Econometrics eJournal*. https://api.semanticscholar.org/ CorpusID:7608876
3. Liu, J. (2024). Navigating the financial landscape: The power and limitations of the ARIMA model. *Highlights in Science, Engineering and Technology*. https://api.semanticscholar.org/CorpusID:270499508
4. Puglia, M., & Tucker, A. (2020). Machine learning, the treasury yield curve and recession forecasting. *Finance and Economics Discussion Series, 2020*. https://doi. org/10.17016/FEDS.2020.038
5. Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785
6. Qin, R. (2022). The construction of corporate financial management risk model based on xgboost algorithm. *Journal of Mathematics*, *2022* (1). https://doi.org/10. 1155/2022/2043369
7. Martin, D., & Póczos, B. (2018). Machine learning-aided modeling of fixed income instruments. https://api.semanticscholar.org/CorpusID:85507134
8. Cheng, W., Lorry, W., & Lin, C.-h. (2001). Forecasting the 30-year us treasury bond with a system of neural networks. *Journal of Computational Intelligence in Finance*, *4*.
9. Sako, K., Mpinda, B. N., & Rodrigues, P. C. (2022). Neural networks for financial time series forecasting. *Entropy*, *24*(5). https://doi.org/10.3390/e24050657
10. Zeng, Z., Kaur, R., Siddagangappa, S., Rahimi, S., Balch, T., & Veloso, M. (2023). Financial time series forecasting using CNN and transformer. https://arxiv.org/abs/ 2304.04912
11. Wu, H., Xu, J., Wang, J., & Long, M. (2022). Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. https://arxiv.org/ abs/2106.13008
12. Zeng, A., Chen, M., Zhang, L., & Xu, Q. (2022). Are transformers effective for time series forecasting? https://arxiv.org/abs/2205.13504